Due Date: May 9, 2004

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE
BEFORE THE BOARD OF PATENT APPEALS AND INTERFERENCES

In re Application of:	
Inventor: Mikael Bisgaard-Bohr et al.	Examiner: Cindy Nguyen
Serial #: 09/739,994	Group Art Unit: 2171
Filed: December 18, 2000	Appeal No.: _____
Title: DATA MODEL FOR ANALYSIS OF RETAIL TRANSACTIONS USING GAUSSIAN MIXTURE MODELS IN A DATA MINING SYSTEM	

BRIEF OF APPELLANTS

MAIL STOP APPEAL BRIEF - PATENTS
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sir:

In accordance with 37 CFR §1.192, Appellants hereby submit the Brief of Appellants, in triplicate, in response to the final rejection received in the above-identified application, as set forth in the Office Action dated December 9, 2003.

Please charge the amount of $330.00 to cover the required fee for filing this Appeal Brief as set forth under 37 CFR §1.17(c) to Deposit Account No. 50-1673 of NCR Corporation, the assignee of the present application. Also, please charge any additional fees or credit any overpayments to Deposit Account No. 50-1673.

I. REAL PARTY IN INTEREST

The real party in interest is NCR Corporation, the assignee of the present application.

II. RELATED APPEALS AND INTERFERENCES

There are no related appeals or interferences for the above-referenced patent application.

III. STATUS OF CLAIMS

Claims 1, 3-9, 11-17, and 19-24 are pending in the application.

Claims 1, 3, 7-9, 11, 15-17, 19, 23, and 24 were rejected under 35 U.S.C. §103(a) as being unpatentable over Fayyad et al., U.S. Patent No. 6,263,337 (Fayyad) in view of Lazarus et al., U.S. Patent No. 6,430,539 (Lazarus).

Claims 4-6, 12-14, and 20-22 were rejected under 35 U.S.C. §103(a) as being unpatentable over Fayyad in view of Lazaraus and further in view of Van Huben et al., U.S. Patent No. 6,327,594 (Van Huben), and these rejections are being appealed.

IV. STATUS OF AMENDMENTS

No amendments to the claims have been made subsequent to the final Office Action.

V. SUMMARY OF THE INVENTION

Appellants' invention, as recited in independent claim 1, is directed to a data structure for analyzing data in a computer-implemented data mining system, wherein the data structure is a data model that comprises a Gaussian Mixture Model that stores transactional data, a basket table that contains summary information about the transactional data, an item table that contains information about individual items referenced in the transactional data, and a department table that contains aggregate information about the transactional data, and the data model is mapped to aggregate the transactional data for cluster analysis.

Appellants' invention, as recited in independent claim 9, is directed to a method for analyzing data in a computer-implemented data mining system, comprising: generating a data structure in the computer-implemented data mining system, wherein the data structure is a data model that comprises a Gaussian Mixture Model that stores transactional data, a basket table that contains summary information about the transactional data, an item table that contains information about individual items referenced in the transactional data, and a department table that contains

aggregate information about the transactional data; and mapping the data model to aggregate the transactional data for cluster analysis.

Appellants' invention, as recited in independent claim 17, is directed to an apparatus for analyzing data in a computer-implemented data mining system, comprising: means for generating a data structure in the computer-implemented data mining system, wherein the data structure is a data model that comprises a Gaussian Mixture Model that stores transactional data, a basket table that contains summary information about the transactional data, an item table that contains information about individual items referenced in the transactional data, and a department table that contains aggregate information about the transactional data; and means for mapping the data model to aggregate the transactional data for cluster analysis.

With regard to the claims, refer to the specification as follows:

(a) at page 3, lines 9-14;

(b) at page 3, line 16 through page 5, line 5, and in FIG. 1 as reference numbers 100-132;

(c) at page 5, line 8 through page 7, line 19;

(d) at page 7, line 22 through page 8, line 8, and in FIG. 2 as reference numbers 200-206; and

(e) at page 8, line 11 through page 9, line 21, and in FIG. 3 as reference numbers 300-310.

VI. ISSUES PRESENTED FOR REVIEW

1. Whether claims 1, 3, 7-9, 11, 15-17, 19, 23, and 24 are obvious under 35 U.S.C. §103(a) in view of Fayyad et al., U.S. Patent No. 6,263,337 (Fayyad) in view of Lazarus et al., U.S. Patent No. 6,430,539 (Lazarus).

2. Whether claims 4-6, 12-14, and 20-22 are obvious under 35 U.S.C. §103(a) as being unpatentable in view of Fayyad, in view of Lazaraus, and further in view of Van Huben et al., U.S. Patent No. 6,327,594 (Van Huben), and these rejections are being appealed..

VII. GROUPING OF CLAIMS

The claims are grouped as follows:

(a) Claims 1, 4, 5, 10, 13, 14, 18, 21 and 22 stand or fall together;

(b) Claims 3, 12 and 20 stand or fall together;

(c) Claims 6, 15 and 23 stand or fall together;

(d) Claims 7, 16 and 24 stand or fall together; and

(e) Claims 8, 17 and 25 stand or fall together.

Separate arguments for each of the groups of claims are provided below.

VIII. ARGUMENTS

A. The Office Action Rejections

In paragraphs (1)-(2) of the Office Action, claims 1, 3, 7-9, 11, 15-17, 19, 23, and 24 were rejected under 35 U.S.C. §103(a) as being unpatentable over Fayyad et al., U.S. Patent No. 6,263,337 (Fayyad) in view of Lazarus et al., U.S. Patent No. 6,430,539 (Lazarus). In paragraph (3) of the Office Action, claims 4-6, 12-14, and 20-22 stand rejected under 35 U.S.C. §103(a) as being unpatentable over Fayyad in view of Lazarus and further in view of Van Huben et al., U.S. Patent No. 6,327,594 (Van Huben).

Appellants' attorney respectfully traverses these rejections.

B. The Appellants' Independent Claims

Independent claim 1 is directed to a data structure for analyzing data in a computer-implemented data mining system, wherein the data structure is a data model that comprises a Gaussian Mixture Model that stores transactional data, a basket table that contains summary information about the transactional data, an item table that contains information about individual items referenced in the transactional data, and a department table that contains aggregate information about the transactional data, and the data model is mapped to aggregate the transactional data for cluster analysis.

Independent claim 9 is directed to a method for analyzing data in a computer-implemented data mining system, comprising: generating a data structure in the computer-implemented data mining system, wherein the data structure is a data model that comprises a Gaussian Mixture Model that stores transactional data, a basket table that contains summary information about the transactional data, an item table that contains information about individual items referenced in the transactional data, and a department table that contains aggregate information about the transactional data; and mapping the data model to aggregate the transactional data for cluster analysis.

Independent claim 17 is directed to an apparatus for analyzing data in a computer-implemented data mining system, comprising: means for generating a data structure in the computer-implemented data mining system, wherein the data structure is a data model that comprises a Gaussian Mixture Model that stores transactional data, a basket table that contains summary information about the transactional data, an item table that contains information about individual items referenced in the transactional data, and a department table that contains aggregate information about the transactional data; and means for mapping the data model to aggregate the transactional data for cluster analysis.

C. The Fayyad Reference

Fayyad describes one exemplary embodiment providing a data mining system for use in finding clusters of data items in a database or any other data storage medium. Before the data evaluation begins a choice is made of the number M of models to be explored, and the number of clusters (K) of clusters within each of the M models. The clusters are used in categorizing the data in the database into K different clusters within each model. An initial set of estimates for a data distribution of each model to be explored is provided. Then a portion of the data in the database is read from a storage medium and brought into a rapid access memory buffer whose size is determined by the user or operating system depending on available memory resources. Data contained in the data buffer is used to update the original model data distributions in each of the K clusters over all M models. Some of the data belonging to a cluster is summarized or compressed and stored as a reduced form of the data representing sufficient statistics of the data. More data is accessed from the database and the models are updated. An updated set of parameters for the clusters is determined from the summarized data (sufficient statistics) and the newly acquired data. Stopping criteria are evaluated to determine if further data should be accessed from the database.

D. The Lazarus Reference

Lazarus describes predictive modeling of consumer financial behavior by application of consumer transaction data to predictive models associated with merchant segments. Merchant segments are derived from consumer transaction data based on co-occurrences of merchants in sequences of transactions. Merchant vectors representing specific merchants are clustered to form

merchant segments in a vector space as a function of the degree to which merchants co-occur more or less frequently than expected. Each merchant segment is trained using consumer transaction data in selected past time periods to predict spending in subsequent time periods for a consumer based on previous spending by the consumer. Consumer profiles describe summary statistics of consumer spending in and across merchant segments. Analysis of consumers associated with a segment identifies selected consumers according to predicted spending in the segment or other criteria, and the targeting of promotional offers specific to the segment and its merchants.

E. The Van Huben Reference

Van Huben describes a common access method to enable disparate pervasive computing devices to interact with centralized data management systems. A modular, scalable data management system is envisioned to further expand the role of the pervasive devices as direct participants in the data management system. This data management system has a plurality of data managers and is provided with a plurality of data managers in one or more layers of a layered architecture. The system performs with a data manager and with a input from a user or pervasive computing device via an API a plurality of process on data residing in heterogeneous data repositories of computer system including promotion, check-in, check-out, locking, library searching, setting and viewing process results, tracking aggregations, and managing parts, releases and problem fix data under management control of a virtual control repository having one or more physical heterogeneous repositories. The system provides for storing, accessing, tracking data residing in said one or more data repositories managed by the virtual control repository. DMS applications executing directly within, on or behalf of, the pervasive computing device organize data using the PFVL paradigm. Configurable managers include a query control repository for existence of peer managers and provide logic switches to dynamically interact with peers. A control repository layer provides a common process interface across all managers. A command translator performs the appropriate mapping of generic control repository layer calls to the required function for the underlying storage engine.

F. Appellants' Independent Claims Are Patentable Over The References

Appellants' invention, as recited in independent claims 1, 9 and 17, is patentable over the references, because the claims recite limitations not found in the references. Specifically, the combination of Fayyad, Lazarus and Van Huben does not disclose a data model that comprises a Gaussian Mixture Model that stores transactional data, a basket table that contains summary information about the transactional data, an item table that contains information about individual items referenced in the transactional data, and a department table that contains aggregate information about the transactional data, and the data model is mapped to aggregate the transactional data for cluster analysis.

The Examiner cites Fayyad as teaching most of the elements of the independent claims, including a data structure for analyzing data in a computer-implemented data mining system, as reference number 12 in FIG. 2 and in the accompanying text. The Examiner also cites Fayyad as teaching that the data structure is a data model that comprises a Gaussian Mixture Model that stores transactional data, at col. 9, lines 22-67. In addition, the Examiner cites Fayyad as teaching that the data model is mapped to aggregate the transactional data for cluster analysis, at col. 8, lines 34-46. However, the Examiner admits that Fayyad does not disclose a basket table that contains summary information about the transactional data, an item table that contains information about individual items referenced in the transactional data, and a department table that contains aggregate information about the transactional data. Nonetheless, the Examiner asserts that Lazarus teaches these elements. Specifically, the Examiner asserts that Lazarus teaches a basket table that contains summary information about the transactional data at col. 13, Table 1, an item table that contains information about individual items referenced in the transactional data at col. 16, lines 13-21, and a department table that contains aggregate information about the transactional data at col. 12, lines 50-63.

Appellants' attorney disagrees. At the locations indicated above, Fayyad and Lazarus, taken individually or in combination, do not teach the claim limitations directed to a data model comprising a Gaussian Mixture Model that stores transactional data, a basket table that contains summary information about the transactional data, an item table that contains information about individual items referenced in the transactional data, and a department table that contains aggregate

information about the transactional data, and the data model is mapped to aggregate the transactional data for cluster analysis.

For example, consider the teaching of Lazarus at col. 13, Table 1:

Lazarus: col. 13, Table 1

TABLE 1
Customer Summary File

Description	Sample Format
Account_id	Char[max 24]
Pop_id	Char ('1'-'N')
Account number	Char[max 16]
Credit bureau score	Short int as string
Internal credit risk score	Short int as string
Ytd purchases	Int as string
Ytd_cash adv	Int as string
Ytd_int purchases	Int as string
Ytd cash adv	int Int as string
State code	Char[max 2]
Zip_code	Char[max 5]
Demographic 1	Int as string
. . .	
Demographic N	Int as string

Nothing in the above discussion of Lazarus teaches "a basket table that contains summary information about the transactional data." Instead, the above discussion relates only to customer summary data.

In another example, consider the teaching of Lazarus at col. 16, lines 13-21:

Lazarus: col. 16, lines 13-21

In order to obtain the initial merchant vectors, additional processing of the master files 408 precedes the analysis of which merchants co-occur in the master files 408. There are two, sequential, processes that are used on the merchant descriptions, stemming and equivalencing. These operations normalize variations of individual merchants names to a single common merchant name to allow for consistent identification of transaction at the merchant. This processing is managed by the vector generation module 510.

Nothing in the above discussion of Lazarus teaches "an item table that contains information about individual items referenced in the transactional data." Appellants' attorney notes that the term "items" is defined in this application as "items purchased by customers," not merchant names. Instead, the above discussion relates only to merchant names.

In yet another example, consider the teaching of Lazarus at col. 12, lines 50-62:

> Lazarus: col. 12, lines 50-62
>
> Customer summary file 404: The customer summary file 404 contains one record for each customer that is profiled by the system, and includes account information of the customer's account, and optionally includes demographic information about the customer. The consumer summary file 404 is typically one that a financial institution, such as a bank, credit card issuer, department store, and the like maintains on each consumer. The customer or the financial institution may supply the additional demographic fields which are deemed to be of informational or of predictive value. Examples of demographic fields include age, gender and income; other demographic fields may be provided, as desired by the financial institution.

Nothing in the above discussion of Lazarus teaches "a department table that contains aggregate information about the transactional data." Instead, the above discussion relates only to customer summary data.

Consequently, the Fayyad and Lazarus Huben references, taken individually or in combination, do not describe a data model comprising a Gaussian Mixture Model that stores transactional data, a basket table that contains summary information about the transactional data, an item table that contains information about individual items referenced in the transactional data, and a department table that contains aggregate information about the transactional data, and the data model is mapped to aggregate the transactional data for cluster analysis.

Moreover, Van Huben fails to overcome these limitations of Fayyad and Lazarus. Recall that Van Huben was only cited for teaching a relational database management system for storing the data model, and then only against the dependent claims.

Thus, the references do not teach or suggest Appellants' invention. Moreover, the various elements of Appellants' claimed invention together provide operational advantages over the references. In addition, Appellants' invention solves problems not recognized by the references.

Thus, Appellants' attorney submits that independent claims 1, 9 and 17 are allowable over the references.

G. Appellants' Dependent Claims Are Patentable Over The References

Dependent claims 3-8, 11-16 and 19-24 are submitted to be allowable over the references in the same manner as the independent claims, because they are dependent on independent claims 1, 9 and 17, respectively, and thus contain all the limitations of the independent claims. In addition, dependent claims 3-8, 11-16 and 19-24 recite additional novel elements not shown by the references.

With regard to claims 3, 11 and 19, which recite that the cluster analysis groups the transactional data into coherent groups according to perceived similarities in the transactional data, the Examiner states that Fayyad teaches these limitations at col. 8, lines 35-64. Appellants' attorney disagrees. Instead, Appellants' attorney submits that Fayyad, at the indicated location, merely describes data structures generally, and vectors of elements specifically, but says nothing about grouping transactional data into coherent groups according to perceived similarities in the transactional data.

With regard to claims 4, 12 and 20, which recite that the data model is stored in a relational database managed by a relational database management system, these claims stand or fall with the independent claims 1, 10 and 18, respectively.

With regard to claims 5, 13 and 21, which recite that the data model is accessed from a relational database managed by a relational database management system, these claims stand or fall with the independent claims 1, 10 and 18, respectively.

With regard to claims 6, 14 and 22, which recite that the data model is mapped into a single flat table format to produce a correct level of aggregation for statistical analysis, the Examiner states that Van Huben teaches these limitations at col. 9, line 66 – col. 10, line 22. Appellants' attorney disagrees. Instead, Appellants' attorney submits that Van Huben, at the indicated location, merely describes the physical implementation of the data management system, but says nothing about mapping a data model into a single flat table format to produce a correct level of aggregation for statistical analysis.

With regard to claims 7, 15 and 23, which recite that the data model is mapped into a database view to produce a correct level of aggregation for statistical analysis, the Examiner states that Fayyad teaches these limitations at col. 8, lines 34-44. Appellants' attorney disagrees. Instead, Appellants' attorney submits that Fayyad, at the indicated location, merely describes data structures

generally, and vectors of elements specifically, but says nothing about database views or mapping a data model into a database view to produce a correct level of aggregation for statistical analysis.

With regard to claims 8, 16 and 24, which recite that the data model is comprised of one row per transaction in the transactional data, the Examiner states that Lazarus teaches these limitations at col. 36, lines 57-65. Appellants' attorney disagrees. Instead, Appellants' attorney submits that Lazarus, at the indicated location, merely describes row descriptions, where each row contains rows for spending and rate in Cash Advances, Purchases, Debits and Total Spending, but says nothing about data models comprised of one row per transaction in the transactional data.

IX. CONCLUSION

In light of the above arguments, Appellants respectfully submit that the cited references do not anticipate nor render obvious the claimed invention. More specifically, Appellants' claims recite novel physical features which patentably distinguish over any and all references under 35 U.S.C. §§ 102 and 103. As a result, a decision by the Board of Patent Appeals and Interferences reversing the Examiner and directing allowance of the pending claims in the subject application is respectfully solicited.

Respectfully submitted,

GATES & COOPER LLP
Attorneys for Appellants

Howard Hughes Center
6701 Center Drive West, Suite 1050
Los Angeles, California 90045
(310) 641-8797

Date: May 5, 2004

By: _____
Name: George H. Gates
Reg. No.: 33,500

GHG/io
G&C 30145.408-US-01

APPENDIX

1. A data structure for analyzing data in a computer-implemented data mining system, wherein the data structure is a data model that comprises a Gaussian Mixture Model that stores transactional data, a basket table that contains summary information about the transactional data, an item table that contains information about individual items referenced in the transactional data, and a department table that contains aggregate information about the transactional data, and the data model is mapped to aggregate the transactional data for cluster analysis.

3. The data structure of claim 1, wherein the cluster analysis groups the transactional data into coherent groups according to perceived similarities in the transactional data.

4. The data structure of claim 1, wherein the data model is stored in a relational database managed by a relational database management system.

5. The data structure of claim 1, wherein the data model is accessed from a relational database managed by a relational database management system.

6. The data structure of claim 1, wherein the data model is mapped into a single flat table format to produce a correct level of aggregation for statistical analysis.

7. The data structure of claim 1, wherein the data model is mapped into a database view to produce a correct level of aggregation for statistical analysis.

8. The data structure of claim 1, wherein the data model is comprised of one row per transaction in the transactional data.

9. A method for analyzing data in a computer-implemented data mining system, comprising:

generating a data structure in the computer-implemented data mining system, wherein the data structure is a data model that comprises a Gaussian Mixture Model that stores transactional data, a basket table that contains summary information about the transactional data, an item table

that contains information about individual items referenced in the transactional data, and a department table that contains aggregate information about the transactional data; and

mapping the data model to aggregate the transactional data for cluster analysis.

11. The method of claim 9, wherein the cluster analysis groups the transactional data into coherent groups according to perceived similarities in the transactional data.

12. The method of claim 9, wherein the data model is stored in a relational database managed by a relational database management system.

13. The method of claim 9, wherein the data model is accessed from a relational database managed by a relational database management system.

14. The method of claim 9, wherein the mapping step comprises mapping the data model into a single flat table format to produce a correct level of aggregation for statistical analysis.

15. The method of claim 9, wherein the mapping step comprises mapping the data model into a database view to produce a correct level of aggregation for statistical analysis.

16. The method of claim 9, wherein the data model is comprised of one row per transaction in the transactional data.

17. An apparatus for analyzing data in a computer-implemented data mining system, comprising:

means for generating a data structure in the computer-implemented data mining system, wherein the data structure is a data model that comprises a Gaussian Mixture Model that stores transactional data, a basket table that contains summary information about the transactional data, an item table that contains information about individual items referenced in the transactional data, and a department table that contains aggregate information about the transactional data; and

means for mapping the data model to aggregate the transactional data for cluster analysis.

19. The apparatus of claim 17, wherein the cluster analysis groups the transactional data into coherent groups according to perceived similarities in the transactional data.

20. The apparatus of claim 17, wherein the data model is stored in a relational database managed by a relational database management system.

21. The apparatus of claim 17, wherein the data model is accessed from a relational database managed by a relational database management system.

22. The apparatus of claim 17, wherein the means for mapping comprises means for mapping the data model into a single flat table format to produce a correct level of aggregation for statistical analysis.

23. The apparatus of claim 17, wherein the means for mapping comprises means for mapping the data model into a database view to produce a correct level of aggregation for statistical analysis.

24. The apparatus of claim 17, wherein the data model is comprised of one row per transaction in the transactional data.

Gates & Cooper LLP
Howard Hughes Center
6701 Center Drive West, Suite 1050
Los Angeles, California 90045

FAX TRANSMISSION TO USPTO

TO: Commissioner for Patents
Attn: Examiner Cindy Nguyen
Patent Examining Corps
Facsimile Center
Alexandria, VA 22313-1450

FROM: George H. Gates
OUR REF.: 9684
TELEPHONE: (310) 642-4146

Total pages, including cover letter: 45

PTO FAX NUMBER: 703/872-9306

If you do NOT receive all of the pages, please telephone us at (310) 641-8797, or fax us at (310) 641-8798.

Title of Document Transmitted:	**BRIEF OF APPELLANTS (IN TRIPLICATE). Charge in the amount of $330.00 to Deposit Account No. 50-1673 of NCR Corporation for filing the Appeal Brief.**
Applicant:	Mikael Bisgaard-Bohr et al.
Serial No.:	09/739,994
Filed:	December 18, 2000
Group Art Unit:	2171
For:	DATA MODEL FOR ANALYSIS OF RETAIL TRANSACTIONS USING GAUSSIAN MIXTURE MODELS IN A DATA MINING SYSTEM
Our Ref. No.:	9684

Please charge all fees to Deposit Account No. 50-1673 of NCR Corporation, the assignee of the present application.

By: _____
Name: George H. Gates
Reg. No.: 33,500

I hereby certify that this paper is being transmitted by facsimile to the U.S. Patent and Trademark Office on the date shown below.

Isabell Ogata
Signature

May 5 2004
Date

G&C 30145.408-US-01

GHG/io